A generalized likelihood ratio test statistic for Cherenkov telescope data

S. Klepser

IFAE, Edifici Cn., Campus UAB, E-08193 Bellaterra, Spain

Abstract

Astrophysical sources of TeV gamma rays are usually established by Cherenkov telescope observations. These counting type instruments have a field of view of few degrees in diameter and record large numbers of particle air showers via their Cherenkov radiation in the atmosphere. The showers are either induced by gamma rays or diffuse cosmic ray background. The commonly used test statistic to evaluate a possible gamma-ray excess is Li and Ma (1983), Eq. 17, which can be applied to independent on- and off-source observations, or scenarios that can be approximated as such. This formula however is unsuitable if the data are taken in so-called "wobble" mode (pointing to several offset positions around the source), if at the same time the acceptance shape is irregular or even depends on operating parameters such as the pointing direction or telescope multiplicity. To provide a robust test statistic in such cases, this paper explores a possible generalization of the likelihood ratio concept on which the formula of Li and Ma is based. In doing so, the multi-pointing nature of the data and the typically known instrument point spread function are fully exploited to derive a new, semi-numerical test statistic. Due to its flexibility and robustness against systematic uncertainties, it is not only useful for detection purposes, but also for skymapping and source shape fitting. Simplified Monte Carlo simulations are presented to verify the results, and several applications and further generalizations of the concept are discussed.

Keywords: test statistic, TeV astronomy, imaging atmospheric Cherenkov technique, Li & Ma

1. Introduction

The field of very-high-energy (VHE, $> 100\,\text{GeV}$) gamma-ray astronomy is currently in its third generation of instruments, with an advanced future project, CTA [1], already being in its preparatory phase. The currently operated systems H.E.S.S. [2], MAGIC [3] and VERITAS [4] have established more than 100 VHE sources in the sky[1]. The acceptence of these telescope systems covers few degrees in diameter, and declines smoothly with increasing distance from the pointing position. Establishing a gamma-ray signal from an astrophysical source requires to significantly prove a gamma-ray excess over background events that typically remain dominant even after selection cuts. This background is mostly composed of diffuse hadrons, part of which appears almost identical to electromagnetic showers and has to be considered to be irreducible [5]. Besides these "gamma-like hadrons", the irreducible background also contains smaller fractions of diffuse electrons and gamma rays. This irreducible background, and the statistical and systematic uncertainties that come with it, are one of the main limiting factors of TeV astronomy; usually, an observational campaign for a given source either reveals one source or none, and only in few cases, or if the effort of a large scan is undertaken, additional unexpected sources are detected. Therefore, a statistical source detection technique that is both sensitive to weak sources and stable against systematic effects is of crucial importance to the field.

The standard test statistic to evaluate an excess of gamma rays from a given sky direction is Li and Ma [6], Eq. 17, which hereafter will be refered to as S_{LM}. It was established among several alternatives they evaluated in their paper, based on the fact that this likelihood ratio test statistic was the only one that yielded a satisfyingly Gaussian null-hypothesis distribution. The formula they presented was designed to compare the event numbers of an on- and off-source observation (N_{on}, N_{off}), and allows for a scaling factor α between the effective observation times ($t_{\text{on}} = \alpha\, t_{\text{off}}$) to account for unequal exposures.

In modern observation practice, most Cherenkov telescope data are *not* taken in On/Off-mode (see Fig. 1, left), because this strategy implies a lot of observation time dedicated to empty sky regions. Also, it is prone to systematic differences between the on- and off-data caused by instabilities in electronic or atmospheric operating conditions, especially if the off-data could not be scheduled contemporaneously enough with the on-source observation. Therefore, usually the "wobble" technique [7] is applied, in which the data are taken at two or more observa-
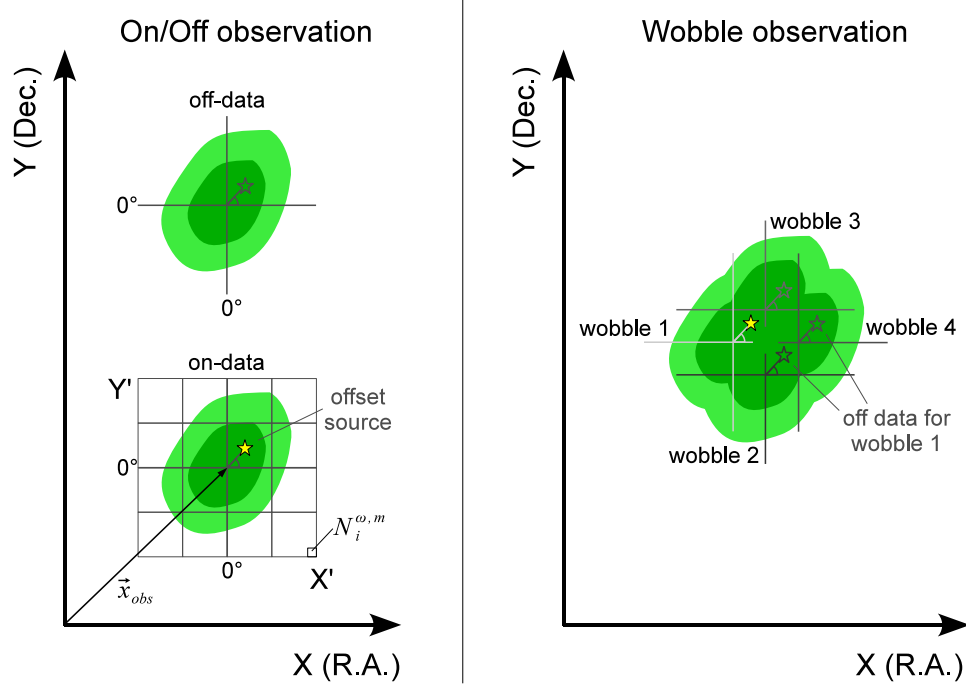
---

[1]http://tevcat.uchicago.edu/

Figure 1: Observation schemes with asymmetric acceptance shapes (green areas). The sky position to be evaluated for a signal is marked with a yellow star, the corresponding off-regions with grey stars. Left: Original On/Off scheme. The off-data provides N_{off}, the on-data N_{on}, and it needs one exposure ratio $\alpha = t_{\text{on}}/t_{\text{off}}$ between those. Right: Wobble scheme with four observation positions. The off data of each wobble set can be taken from all other wobble sets, resulting in four α parameters and their corresponding N_{on}, N_{off}.

tion positions offset in different directions from the main target coordinates (see Fig. 1, right). In this way, each wobble set provides both on- and off-data at the same time. In its original idea, the exposure shape is considered to have some circular symmetry, and for each wobble set the off-data can be taken from the same observation, at sky positions of similar distance to the telescope pointing position ("reflected regions" [8]). In that case, the off-exposure ratio α is the same for all wobble data sets, and S_{LM} can be applied after summing up the on- and off-events of all wobble sets. This constant α can also be achieved approximately if the off-data are taken from *hadron-like* background events ("template background", [9]), or a ring area around the source position ("ring background" [10]), both of which require also symmetry assumptions or an efficiency correction through Monte Carlo simulations of the isotropic background.

In the general case, though, the acceptance symmetry might not hold and a Monte Carlo correction may involve too high uncertainties. This occurs for instance in very low-energy observations, where camera-based acceptance inhomogeneities (dead pixels, trigger fluctuations) are both difficult to model in simulations and furthermore lead to features in the acceptance shape that can depend on the Alt/Az pointing direction or other operation parameters of the system. This is particulary troublesome in two-telescope systems like MAGIC [3], where already the basic geometrical overlap of the two fields of view implies an elongated, Alt/Az-dependent exposure shape. Under these conditions, the "reflected regions" approach does not hold, and the off-data for each wobble set have to be taken from the other wobble sets (see Fig. 1, right). This approach can provide a sensitive measurement, because with several wobble sets, the off regions are numerous and well-populated, but it results in a different α for each wobble set, which is not supported by S_{LM}. On top of that, if this procedure has to be done separately for different types of data (be it for instance different Azimuth angles or telescope multiplicities), it leads to many more α-parameters, and in general, some off-events may happen to be oversampled if they lie in more than one off-region, which is also not considered in S_{LM}. As a consequence, while the background density can still be modeled under certain assumptions and some numerical effort [11], the test statistic S_{LM}, if still applied in some way, becomes very approximative. This may be dealt with in practice by making additional high requirements to the signal-to-background ratio of a detection [3], which however limits the effective sensitivity of the instrument.

Therefore, unlike previous efforts [10, 11], this work will not pursue to extract the variables needed for S_{LM} through the complex task of explicit background modeling. Instead, the likelihood ratio concept behind that test statistic will be generalized, and new formulae will be derived that can directly be applied to multi-wobble Cherenkov telescope data. To do so, no Monte Carlo simulations or exposure symmetry assumptions will be required. It will only be assumed that the telescope acceptance shape is the same for different wobble data sets if they are taken under similar operating conditions[2].

Besides that, this work will also address the disadvantage of S_{LM} that it depends on the size of the signal region in which N_{on} is calculated. This area is usually defined through an integration radius (''θ^2-cut'', with θ being the angular distance between reconstructed gamma direction and source position). Its op-

---

[2]Note that if this basic assumption does not hold, any other symmetry assumption also breaks.

timization either depends on the background density and an assumption of the source strength, or may involve several trials. In this paper, these assumptions are reduced by accomodating the knowledge of the point spread function (PSF) in the formulae, which makes them independent of the source strength or background density.

Likelihood ratios are frequently used to convert a complex likelihood maximization problem to a test statistic that follows a χ^2-distribution. This possibility was first proven in [12], and was suggested for astronomical purposes in [13]. The technique is now widely used in counting type experiments, mainly in X-ray and gamma-ray astronomy, and can be applied both for detection and optimization purposes [14]. It should be pointed out that criticism and potentially more accurate or more general alternatives to the likelihood ratio concept exist [15, 16, 17], but are not subject of this work.

The structure of the paper is to first define the coordinate systems and naming conventions needed for the calculations. Then, a likelihood function is set up and maximized to gain all relevant parameters (Sec. 2). Based on that, the test statistic is formulated in Sec. 3 and its intrinsic inclusion of S_{LM} is demonstrated. In Sec. 4, some further possible generalizations and applications of the formulae are discussed and a recipe for "Likelihood Ratio Skymapping" is suggested. Finally, some example toy simulations are shown in Sec. 5 to verify the method.

2. Building the likelihood function

In this section, a binned likelihood function is formulated which will be the basis of the likelihood ratio test statistic, but can also serve to fit the shape or positional parameters of a detected source. It is set up as a Poissonian probability function that evaluates the consistency of the different wobble subsets taken under a given operating condition with each other, allowing for a hypothetical source with a well-defined shape. It is a binned likelihood, but can be generalized to an unbinned likelihood easily (Sec. 4.5).

In this paper, the term "operating condition" may in practice refer to a range in any quantity that influences the acceptance shape. This binning may for instance be done in the telescope pointing direction (Alt/Az) for two-telescope systems, or in the telescope multiplicity for a multi-telescope system like CTA. It may also be binned in atmospheric conditions, night sky background light level or discrete performance states of the camera. Of course, the formulae are equally valid if no such binning needs to be applied.

2.1. Namings and definitions

The data are assumed to be taken as several wobble sets $\omega = 1, 2, \ldots, W$, centered at sky coordinates $\mathbf{x}_{\text{obs},\omega}$ ("pointing directions"), throughout operating conditions $m = 1, 2, \ldots, M$ (see previous paragraph). Hence, the observation can be described as a set of $W \times M$ two-dimensional sky histograms. These histograms shall be set up in *relative* sky coordinates ($\mathbf{x}' = \mathbf{x} - \mathbf{x}_{\text{obs},\omega}$) centered at the observation direction $\mathbf{x}_{\text{obs},\omega}$ of each wobble set (see grid on the "on-data" drawing of Fig. 1). The individual relative sky bins are named $i = 1, 2, \ldots, I$, where i may be regarded as a representation of a two-dimensional bin (i_x, i_y) without loss of generality. The number of gamma-like events in one such bin (ω, m, i) is $N_i^{\omega,m}$. In relative sky coordinates, the shape of the background event distribution is the same for all histograms that belong to a given operating condition, while a signal at a fixed absolute sky position will appear in different relative locations for different wobble sets. In the following, the axes of the relative sky histograms (X', Y' in Fig. 1) are treated as synonyms for (relative) right ascension and declination, although in practice, one should replace those for axes that are truly rectangular throughout the field of view.

The exposure of the data taken under a given operating condition m is distributed among the wobble sets ω in ratios of a_ω^m, such that $\sum_\omega a_\omega^m = 1$. There are different technical solutions of how to evaluate these ratios, the simplest and most common being through the observation times

$$a_\omega^m = \frac{t_\omega^m}{t_{\text{tot}}^m}. \tag{1}$$

which is correct if the background rate is the same in all wobble sets of m. In the absence of a signal, the expectation value $E(N_i^{\omega,m})$ for a given bin (ω, m, i) equals the background expectation $n_i^{\omega,m}$, which in turn is a well-defined fraction of the total background exposure taken under the operating condition m:

$$E(N_i^{\omega,m}) = n_i^{\omega,m} = a_\omega^m n_i^m. \tag{2}$$

If a signal is present, its shape can be modeled with a (normalized) kernel g_i^ω, which usually may be a function representing the gamma-ray PSF of the instrument, or a more extended shape for dedicated searches of extended sources. The gamma-ray PSF is usually a well-known performance characteristic that is determined from simulations and/or observations of strong known point sources. It might, as any other data cut, slightly depend on the assumed spectral index of the

source. The kernel parameters g_i^ω depend on ω because the relative source coordinates depend on the wobble offset. Also, the shape of the PSF (i.e. the resolution) might depend on the distance from the pointing direction.

The kernel, which only characterizes the shape of the signal, is now multiplied by a scaling factor ϕ to constitute a signal term which is added to the expectation value as follows:

$$E(N_i^{\omega,m} \,|\, \phi) = a_\omega^m \, n_i^m \, (1 \,+\, \phi \, g_i^\omega) \tag{3}$$

This way, the excess events implied by a given *relative excess* ϕ of a source is automatically proportional to the efficiency of the detector at the position i in the operating condition m, and to the exposure of the given wobble set ω. This first-order approximation assumes that the background exposure $n_i^{\omega,m}$, which is proportional to the efficiency to gamma-like hadrons, is also proportional to the gamma-ray efficiency (see Sec. 4.4 for a way to loosen this assumption).

The expectation value of the relative excess ϕ from a gamma-ray source scales with its absolute flux. Therefore, estimators for ϕ can be used for flux skymapping (after the efficiency correction outlined in Sec. 4.4). Still, the meaning of ϕ is technically that of an excess (or deficit), and therefore it is not, like a flux, bound to be positive or zero. Also, its interpretation as a gamma-ray flux is not exclusive — a significantly non-zero excess can also be caused by other physical effects as for instance the moon shadow, or systematic detector artefacts not common to all wobble datasets. In any of these cases, the null hypothesis for the excess/deficit judgement is $\phi = 0$, which is a degenerate case of the signal hypothesis, and not at the edge of the parameter space. This is important for the following, because like this, the likelihood ratio test statistic can be expected to follow a χ^2-distribution. A remaining limitation of the ϕ parameter space is the fact that the expectation value for the event number, $E(N_i^{\omega,m} \,|\, \phi)$, has to be positive, so ϕ has to be larger than $-1/g_{\max}^\omega$.

While the normalization of the kernels g_i^ω has to be consistent among different ω, it neither has to be unity, nor does $\sum_i g_i^\omega$ have to be the same for different ω. For example, if additional off-observations are added to the analysis, they would have $g_i^\omega = 0$ for all i. Also, Sec. 4.4 discusses a possible case where a varying normalization of g_i^ω is appropriate.

For convenience in the following calculations, also the *averaged kernel* for a given operating condition is defined:

$$g_i^m \equiv \sum_\omega a_\omega^m \, g_i^\omega \tag{4}$$

7

and the sum of events in a given bin (m, i):

$$N_i^m \equiv \sum_\omega N_i^{\omega,m} \tag{5}$$

2.2. The likelihood function

The likelihood function can be defined as the product of all Poissonian probability functions of the bins in i, ω and m:

$$L(N_i^{\omega,m} \mid n_i^m, \phi) = \prod_m \prod_i \prod_\omega \frac{(a_\omega^m n_i^m (1 + \phi g_i^\omega))^{N_i^{\omega,m}}}{N_i^{\omega,m}!} e^{-a_\omega^m n_i^m (1 + \phi g_i^\omega)} \tag{6}$$

$$= \prod_m \prod_i e^{-n_i^m (1 + \phi g_i^m)} \prod_\omega \frac{(a_\omega^m n_i^m (1 + \phi g_i^\omega))^{N_i^{\omega,m}}}{N_i^{\omega,m}!} \tag{7}$$

The latter step took advantage of the normalization of a_ω^m within each m and the average kernel convention (Eq. 4).

For the maximization procedure, it is convenient to calculate the log-likelihood $\mathcal{L} = \ln L$, which is

$$\mathcal{L}(N_i^{\omega,m} \mid n_i^m, \phi) = K + \sum_m \sum_i \left[-n_i^m (1 + \phi g_i^m) + \sum_\omega N_i^{\omega,m} \ln (n_i^m (1 + \phi g_i^\omega)) \right] \tag{8}$$

All terms that are independent of the free parameters n_i^m, ϕ were absorbed into the constant K.

2.3. Determination of the parameters

The free parameters to be optimized in order to maximise the likelihood function are $I \times M$ total exposures n_i^m of the operating conditions, and, if a signal at a given sky position is assumed, its corresponding ϕ. The following paragraphs show that this is possible analytically for the n_i^m, and numerically for the relative excess parameter ϕ.

2.3.1. Background density parameters n_i^m

To find the $n_{i,\phi}^m$ that maximize the likelihood function for a given ϕ, one can calculate the first partial derivative of the log-likelihood function (Eq. 8) for a given m' and i':

$$\frac{\partial \mathcal{L}(N_{i'}^{\omega,m} \mid n_i^m, \phi)}{\partial n_{i'}^{m'}} \quad = \quad -1 - \phi g_{i'}^\omega + \sum_\omega \frac{N_{i'}^{\omega,m'}}{n_{i'}^{m'}}$$

$$= \quad -1 - \phi g_{i'}^\omega + \frac{1}{n_{i'}^{m'}} N_{i'}^{m'} \tag{9}$$

8

This expression equals zero if

$$n_{i,\phi}^m = \frac{N_i^m}{1 + \phi\, g_i^m}.$$ (10)

The second derivative of the log-likelihood function is always negative for this solution, so it is the maximum of the likelihood function for a given ϕ. Note also that for the null hypothesis, $\phi = 0$, the best approximator is, intuitively,

$$n_{i,0}^m = N_i^m.$$ (11)

2.3.2. Relative excess parameter ϕ

Inserting the optimized exposures (Eq. 10) into the log-likelihood function (Eq. 8), the log-likelihood expression for a given ϕ parameter can be simplified:

$$\mathcal{L}(N_i^{\omega,m} \,|\, \phi) \;=\; K' + \sum_m \sum_i \left[-N_i^m + \sum_\omega N_i^{\omega,m} \ln\!\left(N_i^m\, \frac{1 + \phi\, g_i^\omega}{1 + \phi\, g_i^m} \right) \right]$$ (12)

$$\;=\; K'' + \sum_m \sum_i \sum_\omega N_i^{\omega,m} \ln\!\left(\frac{1 + \phi\, g_i^\omega}{1 + \phi\, g_i^m} \right)$$ (13)

Again, K' and K'' represent terms that are independent of ϕ.

The partial derivative w.r.t. ϕ is

$$\frac{\partial \mathcal{L}(N_i^{\omega,m} \,|\, \phi)}{\partial \phi} = \sum_m \sum_i \sum_\omega N_i^{\omega,m} \frac{g_i^\omega - g_i^m}{(1 + \phi\, g_i^\omega)(1 + \phi\, g_i^m)}$$ (14)

As it turns out in simulations, finding the root of this expression, and verifying the positive second derivative, is numerically straight-forward. In all reasonable cases, it leads to one solution ϕ_{sup} that maximises the likelihood function (see Sec. 4.2.1 for possible exceptions).

3. The likelihood ratio test statistic

3.1. The likelihood ratio

The likelihood ratio is defined as the ratio of the maximal likelihood of the null hypothesis ($\phi = 0$), and the global maximum of the likelihood, allowing for a

signal ($\phi \neq 0$):

$$
\begin{aligned}
\Lambda &= \frac{L(N_i^{\omega,m} \mid n_{i,0}^m, \phi = 0)}{L(N_{i,\phi}^{\omega,m} \mid n_{i,\phi}^m, \phi)} \\[2mm]
&= \prod_m \prod_i e^{n_{i,\phi}^m (1 + \phi\, g_i^m) - n_{i,0}^m} \prod_\omega \left(\frac{n_{i,0}^m}{n_{i,\phi}^m (1 + \phi\, g_i^\omega)} \right)^{N_i^{\omega,m}} \\[2mm]
&= \prod_m \prod_i \prod_\omega \left(\frac{1 + \phi_{\sup} g_i^m}{1 + \phi_{\sup} g_i^\omega} \right)^{N_i^{\omega,m}}
\end{aligned}
\tag{15}
$$

The typical prescription to convert this to a test statistic [12] is $\mathrm{TS} = -2 \ln \Lambda$. Since the null hypothesis has $M \times I$ parameters, and is a special case of the alternative hypothesis, which has $M \times I + 1$ parameters (because of ϕ), the difference in degrees of freedom is 1, and TS can be expected to follow a χ_1^2 distribution. As in the case of S_{LM}, this only holds for high count numbers; in fact, since the underlying mathematics are the same, it can be assumed that the validity into the low-count regime is similar to that of S_{LM}.

Since $\sqrt{\mathrm{TS}} = \sqrt{\chi_1^2}$ is a half-normal distribution,

$$
S = \sqrt{\mathrm{TS}} = \sqrt{2 \times \sum_m \sum_i \sum_\omega N_i^{\omega,m} \ln\left(\frac{1 + \phi_{\sup} g_i^\omega}{1 + \phi_{\sup} g_i^m} \right)}
\tag{16}
$$

is the Gaussian significance of the considered sky position to contain a gamma-ray excess (or deficit). This holds independently of whether the initial assumptions about the gamma efficiency and PSF are correct, since imprecise assumptions cannot contradict the null hypothesis and might therefore only make the test statistic less sensitive, but never wrong. Section 5.3 shows a validation of this Gaussian nature in simulation.

While the test statistic may be regarded semi-numerical due to the determination of $\phi_{\sup}$, it is unambiguous and can be determined at any desired precision. The trials usually needed to choose the optimal radius of the signal region are obsolete, because the PSF of the instrument is incorpororated, which is typically well-known (see sec. 2.1).

For a given signal $\phi_{\sup}$, the significance after Eq. 16 depends on how much g_i^ω and g_i^m differ. Therefore, a higher number of wobble sets quite naturally leads to a higher significance without increasing the complexity of the analysis by manual selection of valid off-region(s). Furthermore, the formulae are invariant against

10

the number of different operating conditions M, as long as the coverage of the wobble sets is good enough to provide well-balanced exposure ratios a_ω^m needed for the $g_i^{\omega,m}$ to differ.

3.2. The degenerate case of on-off-analysis

In this section, instead of the general multi-wobble case with variable exposure shape and an arbitrary PSF, a case of only one operating condition ($M = 1$) is considered, with separate on- and off-target observations ($W = 2$), and a simple step function kernel that represents a predefined on-target sky region. The observation times are $t_{\text{on}} = \alpha\, t_{\text{off}}$, so the above a_ω relate to the α as follows:

$$a_1 = \frac{\alpha}{\alpha + 1} \tag{17}$$

$$a_2 = \frac{1}{\alpha + 1} \tag{18}$$

Here and in the rest of this section, the index 1 refers to the on-observation and 2 to the off-observation. The step function kernel is 1 within an arbitrarily shaped signal region ($i \in$ SR), and zero elsewhere. Consequently, the kernel constants are

$$g_i^\omega \;=\; \begin{cases} 1 & \text{if } \omega = 1 \text{ and } i \in \text{SR} \\ 0 & \text{else} \end{cases} \tag{19}$$

$$g_i^m \;=\; \begin{cases} \frac{\alpha}{\alpha+1} & \text{if } i \in \text{SR} \\ 0 & \text{else}. \end{cases} \tag{20}$$

Inserting these terms in Eq. 14, it can be converted to

$$\frac{\partial \mathcal{L}(N_i^{\omega,m} \,|\, \phi)}{\partial \phi} \;=\; \frac{\sum_{i \in \text{SR}} N_i^{\omega=1}}{(\alpha + 1)(\phi + 1)(\phi\frac{\alpha}{\alpha+1} + 1)} + \frac{-\alpha\, \sum_{i \in \text{SR}} N_i^{\omega=2}}{(\alpha + 1)(\phi\frac{\alpha}{\alpha+1} + 1)} \tag{21}$$

$$= \; \frac{N_{\text{on}}}{(\alpha + 1)(\phi + 1)(\phi\frac{\alpha}{\alpha+1} + 1)} - \frac{\alpha N_{\text{off}}}{(\alpha + 1)(\phi\frac{\alpha}{\alpha+1} + 1)} \tag{22}$$

the root of which can analytically be calculated to be

$$\phi = \frac{N_{\text{on}} - \alpha N_{\text{off}}}{\alpha N_{\text{off}}}. \tag{23}$$

Putting this to Eq. 16, one finds

$$S = \sqrt{2\left[N_{\text{on}} \ln\left(\frac{1 + \alpha}{\alpha} \frac{N_{\text{on}}}{N_{\text{on}} + N_{\text{off}}} \right) + N_{\text{off}} \ln\left((1 + \alpha) \frac{N_{\text{off}}}{N_{\text{on}} + N_{\text{off}}} \right) \right]} \tag{24}$$

which is the well-known Eq. 17 of Li and Ma [6], making it a degenerate case of Eq. 16 of this work. The advantage of the latter is that it combines an arbitrary amount of differently populated wobble (or off-target) observations, and at the same time uses the actual PSF of the instrument instead of a discrete on-target area. This also avoids the need for an ambiguous choice of integration radius.

4. Further generalizations and suggested applications

4.1. Establishing several sources and their parameters

Once a source is detected, the log-likelihood function (Eq. 13) is fully adequate to estimate the parameters of the source, such as the position or extension. These source parameters can be regarded as parameters of the kernel g_i^ω (which the constant K'' does not depend on).

If one or several sources could be established with relative excesses ϕ_n, and corresponding kernel parameters $g_{i,n}^\omega$, $g_{i,n}^m$, they can be inserted to Eq. 3 in order to scan the field for other sources, and cross-check the new null-hypothesis distribution:

$$E(N_i^{\omega,m} \mid \phi) = a_\omega^m \, n_i^m \, (1 \; + \; \phi \, g_i^\omega \; + \; \sum_n \phi_n \, g_{i,n}^\omega) \tag{25}$$

In this case, the log-likelihood derivative (Eq. 14) turns to

$$\frac{\partial \mathcal{L}(N_i^{\omega,m} \mid \phi)}{\partial \phi} = \sum_m \sum_i \sum_\omega N_i^{\omega,m} \, \frac{g_i^\omega(1 + \sum_n \phi_n \, g_{i,n}^m) - g_i^m(1 + \sum_n \phi_n \, g_{i,n}^\omega)}{(1 + \phi \, g_i^\omega + \sum_n \phi_n \, g_{i,n}^\omega)(1 + \phi \, g_i^m + \sum_n \phi_n \, g_{i,n}^m)}, \tag{26}$$

and also the test statistic derived from the likelihood ratio is now different, because the null hypothesis already assumes the presence of sources:

$$S \; = \; \sqrt{\mathrm{TS}} \; = \; \sqrt{2 \times \sum_m \sum_i \sum_\omega N_i^{\omega,m} \, \ln\left(\frac{1 \; + \; \frac{\phi_{\mathrm{sup}} \, g_i^\omega}{1 + \sum_n \phi_n \, g_{i,n}^\omega}}{1 \; + \; \frac{\phi_{\mathrm{sup}} \, g_i^m}{1 + \sum_n \phi_n \, g_{i,n}^m}}\right)} \tag{27}$$

Using these modified formulae is correct as long as the sources and their effective off-regions are spatially independent of each other. If not, then their averaged kernels $g_{i,n}^m$ overlap, and the ϕ_n cannot be optimized independently, in which case a multi-parameter maximization, or an iterative scheme has to be applied.

12

4.2. Likelihood Ratio Skymapping

A common task with Cherenkov telescope data is to scan the whole field of view for unknown sources. Since the formulae provided here are perfectly viable to do this even in the general case of an unknown, operating-condition-dependent acceptance shape, I suggest the following "Likelihood Ratio Skymapping" procedure:

1. The events are filled into histograms $N_i^{\omega,m}$, and the relative exposure ratios a_ω^m are determined.
2. A grid of absolute sky positions is defined. The bins are independent of those of the wobble histograms and should be large enough to avoid over-sampling (see Sec. 4.2.1).
3. For each grid point, the following procedure is applied:
    (a) The kernel constants g_i^ω are calculated for each wobble set ω in relative sky coordinates.
    (b) The average kernels g_i^m are calculated, using the weights a_ω^m (Eq. 4).
    (c) The root of Eq. 14 is determined to obtain ϕ_{sup}.
    (d) The significance is calculated after Eq. 16, using the sign of ϕ_{sup} to distinguish positive from negative excess.
4. If a source is found, it is modeled with a log-likelihood fit (Eq. 13), and the whole procedure is repeated with the modified formulae of 4.1 until the significance distribution follows the null-hypothesis distribution of a Gaussian function.

In this scheme, which follows a similar iterative strategy as the likelihood analysis of the Fermi Science Tools[3], the final result is a number of established sources, an empty significance skymap, and a significance distribution that meets the null hypothesis. There are several ways one might display these sources in a publication while avoiding the negative excesses elsewhere, but in this work, no fixed recommendation concerning this graphical issue will be pursued.

4.2.1. Technical notes and caveats

Calculating and interpreting a skymap requires some more considerations than the mere calculation of a significance value. Some of these concern features that can be regarded as caveats of the method introduced here, and some are features that equally occur in other VHE skymapping methods.

---

[3]http://fermi.gsfc.nasa.gov/ssc/data/analysis

13

1. Features common to most skymapping methods
   (a) In very poorly populated data sets, it can happen that the log-likelihood derivative (Eq. 14) does not approach zero (in empty sky regions), or the likelihood increases towards negative ϕ until it meets its limit $-1/g_{\max}^\omega$. Both are caveats that equally apply to other skymapping methods, because the negative excess that can be considered is always limited by the fact that no negative event numbers can be measured. In Sec. 5, satisfying results are obtained by discarding skybins that belong to the former case (in areas without sensitivity), and using the limit of S when ϕ approaches $-1/g_{\max}^\omega$.
   (b) If the grid of the skymap is small compared to the kernel/PSF of the instrument, the single significance values S become correlated and their distribution cannot be easily tested to be compatible with a Gaussian or not. This oversampling problem is also common to most skymapping methods and needs to be addressed with an appropriate binning or trial strategy (see Sec. 5), or an appropriate consideration of the correlations in the compatibility test.
   (c) In case of multiple sources (or, equivalently, very extended sources), signal photons of one source might appear in the effective off regions of another, weakening the significances of both. This issue can be addressed by observing in multiple wobble sets, and a sufficient sky coverage around a potentially extended source.
2. Features specific to Likelihood Ratio Skymapping:
   (a) Since fewer assumptions go into the test statistic than in conventional background estimation techniques [10], it is much less affected by systematic uncertainties. However, it is also bound to be statistically somewhat less sensitive, at least in the two-wobble case where each of the two data sets only contributes one off region. If many wobble coordinates are used, every data set provides $W-1$ off regions, and the reduced significance can be expected to be compensated rapidly.
   (b) As S_{LM}, also the formulae presented here do not distinguish between excess and deficit of events. Therefore, the presence of a signal produces a positive S peak at the source position, but also a negative S peak at the sky coordinates that the corresponding off-data are taken from. This effect cannot provoke false detections, and is compensated by the modified formulae in Sec. 4.1. Also, it looses importance with increasing number of wobble sets.

4.3. Excess events

If a number of excess events has to be calculated, this can be done summing up the signal part of Eq. 3:

$$N_{\text{ex}} \;=\; \sum_m \sum_i \sum_\omega a_\omega^m \, n_i^m \, \phi_{\text{sup}} \, g_i^\omega \tag{28}$$

$$=\; \sum_m \sum_i N_i^m \, \frac{\phi_{\text{sup}} \, g_i^m}{1 + \phi_{\text{sup}} \, g_i^m} \tag{29}$$

In case the signal of other sources has to be considered, as discussed in Sec. 4.1, the corresponding expression is

$$N_{\text{ex}} \;=\; \sum_m \sum_i N_i^m \, \frac{\phi_{\text{sup}} \, g_i^m}{1 + \phi_{\text{sup}} \, g_i^m + \sum_n \phi_n \, g_{i,n}^m} \,. \tag{30}$$

4.4. Variable gamma/hadron acceptance ratio, flux skymapping and variable sources

The relative excess parameter ϕ is based on the assumption that the gamma-ray acceptance is proportional to the acceptance for gamma-like hadrons. However, Monte Carlo simulations may indicate that this is not the case, and provide correction factors

$$\gamma_i^m \;=\; \frac{\varepsilon_i^m(\text{gammas})}{\varepsilon_i^m(\text{hadrons})}, \tag{31}$$

where the ε are the efficiencies to gammas and hadrons. In the scheme presented here, these constants may simply be incorporated into the kernel constants:

$$g_i^\omega \;\rightarrow\; \gamma_i^m \, g_i^\omega \tag{32}$$

This way, or if $\gamma_i^m = 1$ is a good approximation, the factor between hypothetical source flux and relative excess parameters ϕ_{sup} will be constant throughout the field of view, and ϕ_{sup} may be used to display the physically more relevant *relative flux* skymap (possibly omitting unphysical negative flux values for consistency). One has to bear in mind though that ϕ_{sup} is a value that is relative to the background density, so it is only reliable if the background density is sufficiently high. In poorly populated skymaps, the number of excess events (Eq. 29) may be a more stable parameter.

Another considerable case where the kernel normalization may be modified is when the signal is not constant in time, but variable, and an a-priori assumption of the light curve is available.

4.5. Unbinned analysis and "orbit" mode

The presented formulae can be adopted for unbinned analysis if the limit of an infinitely fine sky-binning I is assumed. In this case, most bins are empty ($N_i^{\omega,m} = 0$) and the sums in Equations 13, 14 and 16 turn into sums over the relative coordinates $\mathbf{x_j'}$ of the single events $j = 1, 2, \ldots, J$ (i.e. bins for which $N_i^{\omega,m} = 1$). Instead of the previously discrete kernel constants, functions $g^\omega(\mathbf{x_j'})$ have to be used. In this case, the log-likelihood function is

$$\mathcal{L}(\mathbf{x_j'} \,|\, \phi) = K'' + \sum_j \ln\left(\frac{1 + \phi\, g^{\omega_j}(\mathbf{x_j'})}{1 + \phi\, g^{m_j}(\mathbf{x_j'})} \right), \tag{33}$$

the derivative w.r.t ϕ is

$$\frac{\partial \mathcal{L}(\mathbf{x_j'} \,|\, \phi)}{\partial \phi} = \sum_j \frac{g^{\omega_j}(\mathbf{x_j'}) - g^{m_j}(\mathbf{x_j'})}{(1 + \phi\, g^{\omega_j}(\mathbf{x_j'}))(1 + \phi\, g^{m_j}(\mathbf{x_j'}))}, \tag{34}$$

and the test statistic is

$$S = \sqrt{2 \times \sum_j \ln\left(\frac{1 + \phi_{\sup}\, g^{\omega_j}(\mathbf{x_j'})}{1 + \phi_{\sup}\, g^{m_j}(\mathbf{x_j'})} \right)}. \tag{35}$$

Although potentially more precise, this unbinned quantities may be computationally much more expensive, since the function $g^\omega(\mathbf{x'})$ has to be calculated $J \times W$ times for each evaluation of $\mathcal{L}$, whereas for a binned analysis, only $I \times W$ calculations are needed, which are typically much fewer calls.

A case where the unbinned approach might be computationally more effective than the binned approach is the application to "orbit mode" observations [18], in which the telescope pointing is not discrete, but moves along a circle around the target coordinates. If one introduces the *wobble angle* ξ that parametrizes the pointing direction along that circle, the above equations still hold if $g^{\omega_j}(\mathbf{x_j'})$ is replaced by a variable kernel $g(\mathbf{x_j'}, \xi)$, and $g^{m_j}(\mathbf{x_j'})$ is the convolution of $g(\mathbf{x_j'}, \xi)$ with the distribution of ξ within the operating condition m_j that the event j belongs to.

5. Monte Carlo Simulations

To verify the proposed Likelihood Ratio Skymapping scheme and its practical feasibility, two standard scenarios were simulated for which the method might be useful. For simplicity, again arbitrary X-Y-coordinates are used instead of right

ascension and declination. The histograms $N_i^{\omega,m}$ are $3° \times 3°$ and the kernel (PSF) is assumed to be a Gaussian function with $\sigma_{39\%} = 0.05°$. The bin sizes are $1\,\sigma_{39\%}$ (although they may in principle be arbitrarily small, at expense of computational load). The sky grid is conservatively chosen to be very coarse ($\sqrt{2\pi}\,\sigma_{39\%})^4$, to minimize the correlations between the significances. In practice, a finer grid may be chosen if the oversampling correlations are taken into account, which is not pursued here for clarity. The exposure ratios a_ω^m are considered to be on-time ratios, and the total number of generated background events are 80000. For each of the simulations, the following plots are provided, both for in the presence and absence of a source in the field of view:

- Measured distributions of events ($N_i^{\omega,m}$).

- Kernels g_i^ω and g_i^m (Eq. 4), assuming the coordinates of the source.

- Likelihood as a function of ϕ, assuming again the coordinates of the source, to show how well ϕ_{sup} is defined.

- Resulting significance skymap and distribution.

The root-finding was done with the bisection method, and no other approximations were made, although the fact that most summands in Eq. 14 are almost zero may easily be exploited to build a more efficient algorithm.

5.1. Case 1: W=2, M=1, asymmetric exposure shape, equal on-time, source at observation center

This is, for an asymmetric exposure, a simple, ideal case. The source is at (0/0) in absolute sky coordinates, and the two wobble positions are offset by $\pm 0.4°$ in X-direction. Figure 2 shows the results. No significant excess can be found and the null-hypothesis distribution is compatible with a Gaussian with a mean of zero and with a width of $\sigma = 1$.

Despite the asymmetric exposure, S_{LM} can be evaluated correctly, summing up the off-events for a given on-region in the same region of the opposite wobble

---

[4]The effective sky area covered by a two-dimensional Gaussian kernel is $2\pi\sigma^2$. In order for this to equal the area of a rectangular grid cell, and thus lead to the same number of trials as if the kernel had the shape of a grid cell, the grid constant has to be $\sqrt{2\pi}\,\sigma$

set. As mentioned in 3.2, the radius inside which events are integrated is basically a free parameter[5], so the significance is scanned in a range of radii between $1 - 3\,\sigma_{\mathrm{PSF}}$, to make sure that the comparison is not artificially biased in the favour of the new formulae presented here. In the absence of a signal (Fig. 2), the significance found at (0/0) is $0.4\,\sigma$ with Eq. 16 and $S_{\mathrm{LM}} = 1.1\,\sigma$.

Figure 3 is the same scenario, but inserting 300 excess events at the observation center. The excess is detected at $6.6\,\sigma$ with Eq. 16 and $S_{\mathrm{LM}} = 6.3\,\sigma$. Since the source happens to be exactly on a grid point (at (0/0)), basically only one grid point shows the signal, proving that the correlations between the significances are negligible. In the significance distribution, this one value peaks out while not disturbing the general Gaussian appearance of the bulk of entries. As negative ϕ are allowed (in order to see the negative part of the distribution), the ambiguity of excess and lacking acceptance causes minor downward artefacts at $(0.0/\pm 0.8)$ in the skymap and at the negative end of the significance distribution. This is an expected feature of the algorithm (see Sec. 4.2.1).

Finally, Fig. 4 is the same setup as before, but the detected source is included in the null hypothesis to validate the detection and look for possible other sources (see Sec. 4.1). The significance map is again empty, the downward artefacts are gone and the new null hypothesis is met.

5.2. Case 2: W=4, M=2, pointing-dependent exposure shape, very different exposure times, one additional off-data set, extended source with large offset

This is a very complex scenario that benchmarks the flexibility of the test statistic and the Likelihood Ratio Skymapping technique. The background events are distributed over 7 sub-histograms consisting of 3 wobble sets taken in two different operating conditions, and an additional set of off-data for one of the operating conditions. As can be seen in Fig. 5, the data can be combined without any problem, yielding an empty skymap and a Gaussian null-hypothesis distribution.

Figure 6 is the same scenario, but with 1200 photons added to mimic an extended source ($\sigma_{\mathrm{src}} = 0.2°$) at the large-offset sky coordinate (0.4/1.0). Even using the same PSF-Kernel as before, the source can be detected ($5.5\,\sigma$).

5.3. Gaussian property of Eq. 16 and comparison to S_{LM}

Figure 7 shows a simulation of significances calculated for the whole field of view of 20 signal-free skymaps simulated as in case 2 (Sec. 5.2). The distri-

---

[5]There are recipes to determine which integration radius is appropriate, but a discussion or comparison of those is not within scope of this paper.
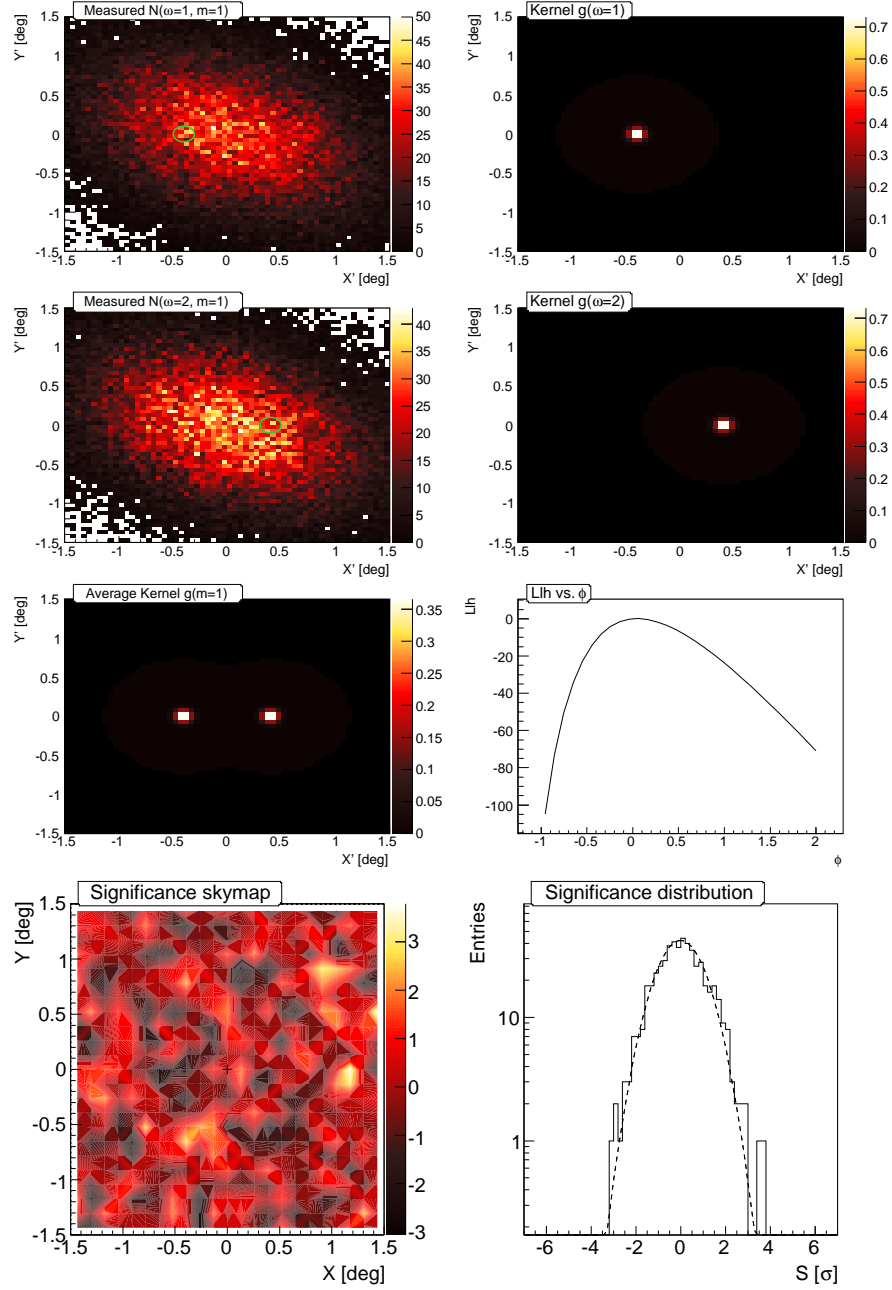
Figure 2: Case 1: W=2, M=1, asymmetric exposure shape, equal on-time, no source. In the upper left, the simulated measured distributions are shown, in coordinates relative to the observation center. Right of it are the kernels g_i^ω, assuming the source to be at the observation center, and below is the corresponding average kernel g_i^m. The graph in the third row on the right is the log-likelihood as a function of ϕ, and in the bottom row, the skymap and significance distribution are shown (the dotted line is a Gaussian of $\mu = 0$ and $\sigma = 1$). The observation center is marked with a cross in the sky map, and green circles in the data histograms.
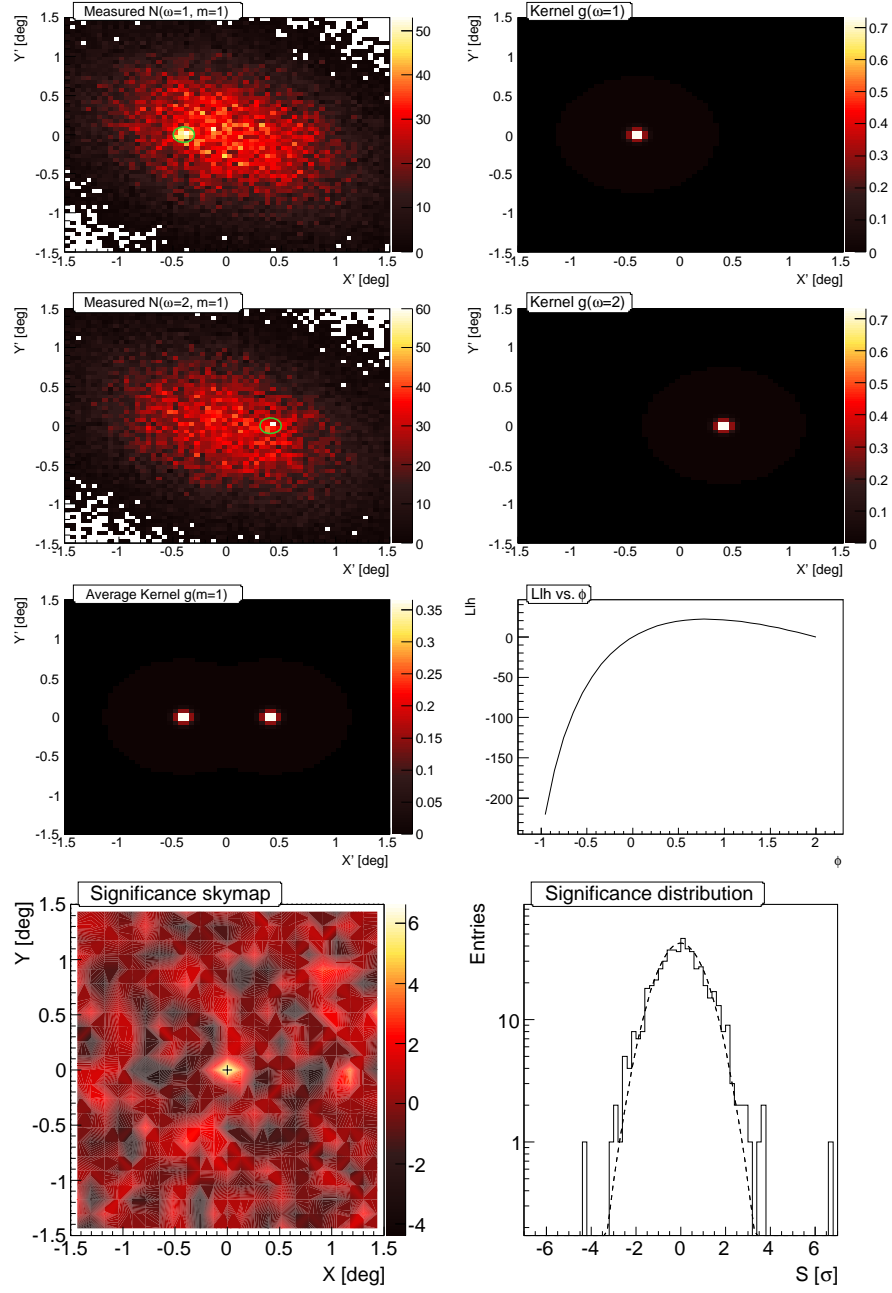
Figure 3: Same as Fig. 2, but with 300 excess events introduced at the observation center.
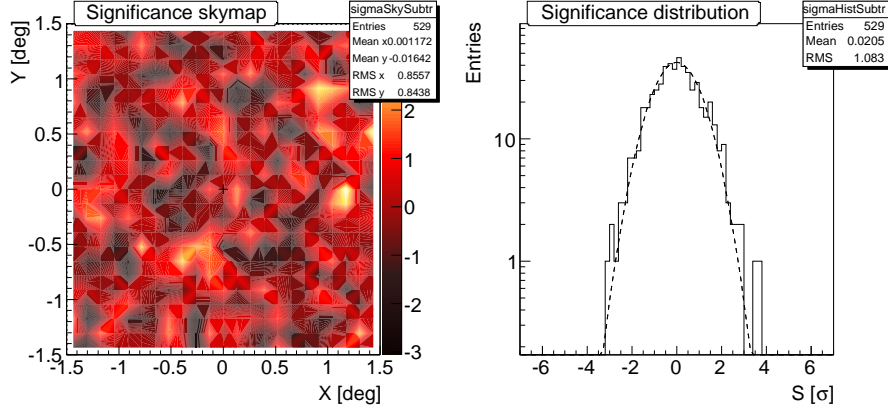
Figure 4: Same as Fig. 3, but with the detected excess included as an established source in the null hypothesis. The significance map is again consistent with the null hypothesis. Consequently, the log-likelihood function peaks at a relative excess parameter (ϕ) that is significantly larger than zero.

bution contains also bins of comparably small exposure and bins for which the optimization of the ϕ-parameter reached its boundary condition. No treatment of the remaining correlation between the significances is undertaken. A Poissonian likelihood fit to the distribution yields Gaussian parameters of $\mu = -0.012 \pm 0.010$ and $\sigma = 1.006 \pm 0.007$. The $\chi^2 = 40.5$ may be compared to the 37 non-zero bins of the distribution. This verifies both Eq. 16 as a Gaussian significance and the suggested recipes to deal with its features in the skymapping case.

In the above case 1, the test statistic of Eq. 16 is somewhat higher than S_{LM}, even though the latter involves an optimization of the applied integration radius (which is expected to lead to a selection bias). To study this possible trend systematically, Figure 8 shows the distribution of many significances obtained for the same case simulated 1000 times. It shows that while Eq. 16 does not always lead to a higher significance, there is a significant trend that in average it does.

Repeating this study with a lower background level, however, the effect is much weaker, while with a higher background density, the effect is stronger. This is assumably because with a low background density, the optimized θ^2-cut will be large and include the whole signal, reducing the advantage of the known kernel from which Eq. 16 benefits.

The conclusion can be drawn that the test statistic presented in this work performs slightly better in cases with low signal-to-background ratios, and similar in cases where the dominant statistical uncertainty is that of the signal itself.
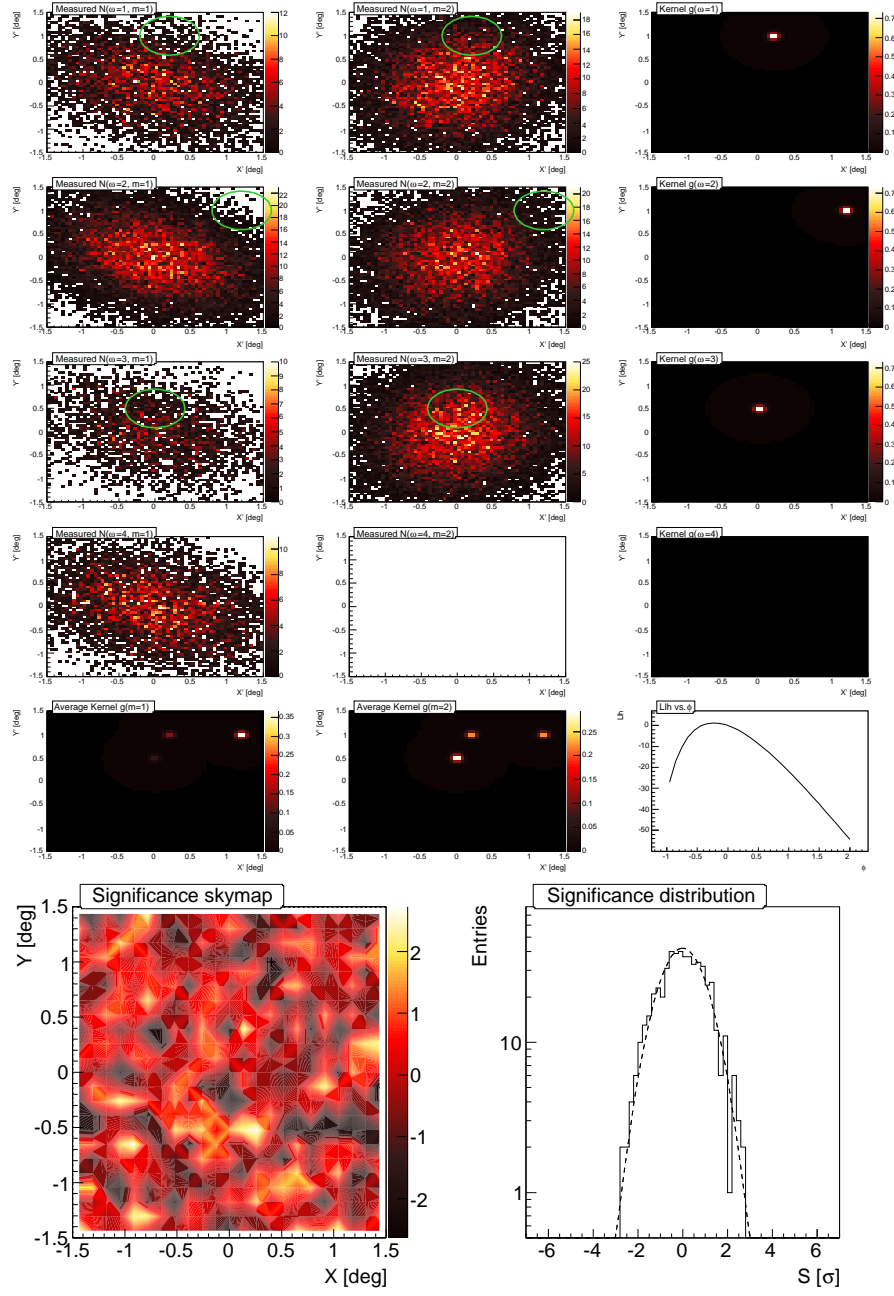
Figure 5: Case 2: W=4, M=2, pointing-dependent exposure shape, very different exposure times, one additional off-data set, The example kernel plots refer to the sky position (0.4/1.0), marked with a cross in the significance skymap, where a source is injected in Fig. 6.
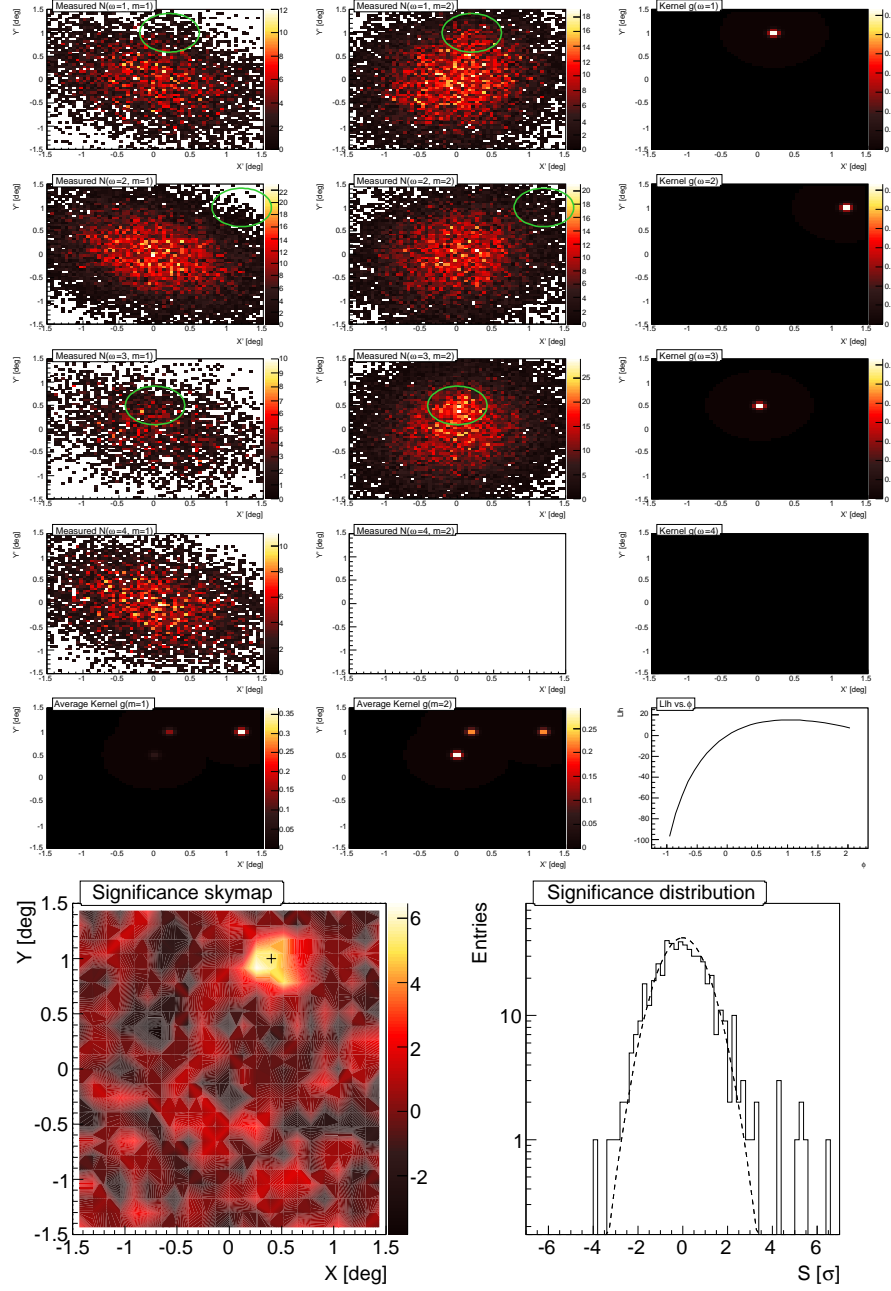
Figure 6: Same as Fig. 5, but with 1200 excess events introduced at the position marked with a green circle in the upper histograms.
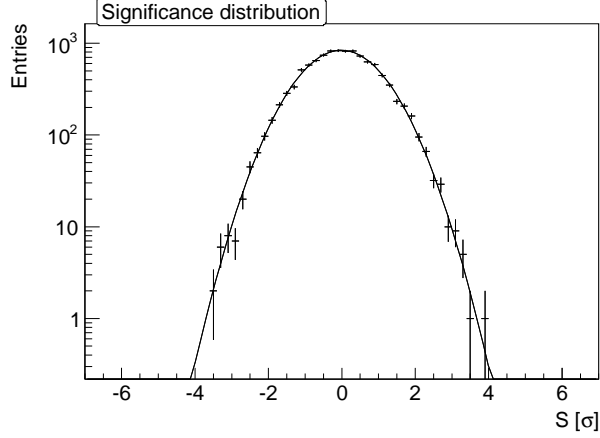
Figure 7: Accumulated significance distribution of 20 simulated skymaps as in Sec. 5.2. Despite the complex setup of several wobble positions and operating conditions, and the skymapping caveats outlined in Sec. 4.2.1, the significance distribution follows very well the shape of a Gaussian distribution (the solid line is the Gaussian fit to the data discussed in the text).
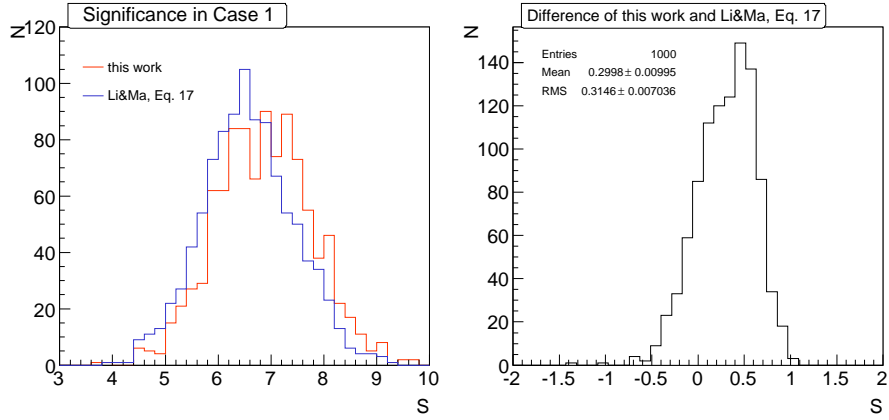

Figure 8: Comparison of significances obtained with Eq. 16 and S_{LM} (with optimized θ^2-cut). Despite the trials implied by the optimized θ^2-cut, Eq. 16 yields a systematically higher significance.

24

6. Conclusions

This paper derives a new generalized test statistic (Eq. 16) that can be used for significance calculation in VHE astronomy. The advantages over the existing test statistics are that it flexibly takes into account any number of data subsets from different wobble coordinates and operating conditions of the system, even if the acceptance shape is very irregular and different between these operating conditions. Also, it takes advantage of a known gamma-ray PSF while not requiring the optimization of an integration radius ("θ^2-cut").

The test statistic can be applied to any position in the field of view, so it is very suitable for skymapping purposes. The advantages of this approach is that the test statistic only makes minimal assumptions on the acceptance field of view and does not require any exposure symmetry or Monte Carlo simulations. It is hence unaffected by many systematic uncertainties. A "Likelihood Ratio Skymapping" procedure is suggested in Sec. 4.2.

The log-likelihood function (Eq. 13) can also be applied to fit the shape and position parameters of the source. The formulae are furthermore extendable to accomodate established sources in the field of view, a non-homogeneous PSF shape or gamma-to-hadron acceptance ratio, an unbinned analysis approach or the "orbit" observation mode. If the background event density is sufficiently high, the relative excess parameter ϕ is well-suited to calculate a gamma-ray flux map of the field of view.

In several simulated scenarios it is verified that the test statistic can not only handle the difficult situations it is designed for, but also seems to be systematically higher in the signal case, and therefore more sensitive, than the commonly used test statistic after Li and Ma [6], Eq. 17.

7. Acknowledgements

References

[1] CTA Consortium, Design Concepts for the Cherenkov Telescope Array. `arXiv:1008.3703`.

[2] J. A. Hinton, et al., The status of the HESS project, New Astronomy Reviews 48 (2004) 331–337. `arXiv:astro-ph/0403052`, `doi:10.1016/j.newar.2003.12.004`.

[3] J. Aleksić, et al., Performance of the MAGIC stereo system obtained with Crab Nebula data, accepted for publication in Astropart. Phys.`arXiv:1108.1477`.

[4] J. Holder, et al., The first VERITAS telescope, Astroparticle Physics 25 (2006) 391–401. `arXiv:astro-ph/0604119`, `doi:10.1016/j.astropartphys.2006.04.002`.

[5] D. Sobczynska, Natural limit on the γ/hadron separation for a stand alone air Cherenkov telescope, Journal of Physics G Nuclear Physics 34 (2007) 2279–2288. `arXiv:astro-ph/0702562`, `doi:10.1088/0954-3899/34/11/005`.

[6] T.-P. Li, Y.-Q. Ma, Analysis methods for results in gamma-ray astronomy, ApJ 272 (1983) 317–324. `doi:10.1086/161295`.

[7] V. P. Fomin, A. A. Stepanian, R. C. Lamb, D. A. Lewis, M. Punch, T. C. Weekes, New methods of atmospheric Cherenkov imaging for gamma-ray astronomy. I. The false source method, Astropart. Phys. 2 (1994) 137–150. `doi:10.1016/0927-6505(94)90036-1`.

[8] F. Aharonian, et al., Evidence for TeV gamma ray emission from Cassiopeia A, A&A 370 (2001) 112–120. `arXiv:astro-ph/0102391`, `doi:10.1051/0004-6361:20010243`.

[9] G. P. Rowell, A new template background estimate for source searching in TeV gamma -ray astronomy, A&A 410 (2003) 389–396. `arXiv:astro-ph/0310025`, `doi:10.1051/0004-6361:20031194`.

[10] D. Berge, S. Funk, J. Hinton, Background modelling in very-high-energy γ-ray astronomy, A&A 466 (2007) 1219–1229. `arXiv:astro-ph/0610959`, `doi:10.1051/0004-6361:20066674`.

[11] S. Lombardi, K. Berger, P. Colin, et al., Advanced stereoscopic gamma-ray shower analysis with the MAGIC telescopes, in: Proc. ICRC 2011 (arXiv:1109.6195), 2011.

[12] S. S. Wilks, The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypothe
Ann. Math. Statist. 9 (1) (1938) 60–62.
URL http://projecteuclid.org/euclid.aoms/1177732360

[13] W. Cash, Parameter estimation in astronomy through application of the like-
lihood ratio, ApJ 228 (1979) 939–947. doi:10.1086/156922.

[14] J. A. Nousek, D. R. Shue, Chi-squared and C statistic minimization for low
count per bin data, ApJ 342 (1989) 1207–1211. doi:10.1086/167676.

[15] S. Gillessen, H. L. Harney, Significance in gamma-ray astronomy - the
Li & Ma problem in Bayesian statistics, A&A 430 (2005) 355–362.
arXiv:astro-ph/0411660, doi:10.1051/0004-6361:20035839.

[16] R. Protassov, D. A. van Dyk, A. Connors, V. L. Kashyap, A. Siemigi-
nowska, Statistics, Handle with Care: Detecting Multiple Model Com-
ponents with the Likelihood Ratio Test, ApJ 571 (2002) 545–559.
arXiv:astro-ph/0201547, doi:10.1086/339856.

[17] S. N. Zhang, D. Ramsden, Statistical data analysis for gamma-
ray astronomy, Experimental Astronomy 1 (1990) 145–163.
doi:10.1007/BF00462037.

[18] G. Finnegan, for the VERITAS Collaboration, Orbit Mode obser-
vation Technique Developed for VERITAS, Proc. Fermi Symposium
2011,arXiv:1111.0121.